

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)**ScienceDirect**

Procedia - Social and Behavioral Sciences 103 (2013) 77 – 85

**Procedia**  
Social and Behavioral Sciences13<sup>th</sup> International Educational Technology Conference

# A Survey on Teachers' Awareness and Attitudes on Computer-Corpus Data: An Assisted Technology-based EFL Vocabulary Selection and Instruction Source

Fatima Zohra Belkhir \*

*Abou-bekr Belkaid University, Faculty of Letters and Languages, Larocade 13000, Tlemcen, Algeria*

---

## Abstract

The present research paper focuses on the area of vocabulary teaching in the foreign language classroom. It attempts to highlight the way computer corpus data ease the difficulties involved in EFL (English Foreign Language) vocabulary teaching and learning. The paper aims at exploring the extent of awareness of EFL teachers to the concept of computer corpus-data, raising awareness on the efficacy of computer-corpus data in the area of EFL vocabulary selection and instruction and analyzing teachers' attitudes on the application of those data as a language source in selecting EFL vocabulary teaching materials.

© 2013 The Authors. Published by Elsevier Ltd. Open access under [CC BY-NC-ND license](#).

Selection and peer-review under responsibility of The Association of Science, Education and Technology-TASET, Sakarya Universitesi, Turkey.

Keywords: vocabulary; computer-corpus data; EFL teachers; awareness; attitudes

---

## 1. Introduction and statement of the problem

Learning a foreign language involves the development of different skills, namely, reading, writing, listening and speaking. Although specific skills may be focused on separately for pedagogical reasons, two elements are crucial to the process of acquiring and using a language: vocabulary and grammar (Celce-Murcia, 2001). However, the importance of grammar has been widely recognized in English Language Teaching (ELT) methodologies than vocabulary (Celce-Murcia, 2001; Nunan, 2001; Swan, 2002). This is because the latter was incidental to the main purpose of the target language teaching. In this respect, DeCarrico (2001:285) writes:

---

\* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000 .

E-mail address: [fatimazohraonline@yahoo.fr](mailto:fatimazohraonline@yahoo.fr)

*“vocabulary has not always been recognized as a priority in language teaching”*. In addition, Richards and Renandya (2002) affirm that vocabulary teaching and learning has been left to a position of secondary importance if compared with other fields of research in language teaching and learning. However, from the mid 1980s onwards, vocabulary becomes the area of interest of many investigators (Laufer, 1986; Carter, 1987, 1988; Sinclair & Renouf, 1988; Nation, 1990; Willis, 1990; Descamps, 1992; Nattinger & DeCarrico, 1992; Lewis, 1993, 1997; Read, 2000 etc.).

The growth of interest in vocabulary learning pushes many researchers to investigate its effectiveness on successful Second Language (SL) learning. For example, Laufer (1986) and Nation (1990) realize that many of SL learners' receptive skills (listening and reading) and productive skills (speaking and writing) difficulties result from an inadequate vocabulary. They also assert that even when learners are at higher levels of language competence and performance, they are still in need to vocabulary learning for this latter is a lifelong process. On their part, Richards J.C. and Renandya (2002: 252) find out that: *“Vocabulary is a core component of language proficiency and provides much of the basis for how well learners listen, speak, read and write. Without an extensive vocabulary... learners often achieve less than their potential and may be discouraged from making use of language”*. In the same context, Beglar and Hunt (2005: 7) confess that *“Vocabulary acquisition is a crucial, and in some senses, the central component in successful foreign language acquisition”*. Accordingly, it could be argued that vocabulary is the keystone of successful Second Language and Foreign Language (FL) learning and its acquisition is the basic foundation of the target language proficiency since its knowledge pervades all four language skill areas and helps learners in both language comprehension and production. Therefore, words are perhaps the greatest tools teachers can give to their learners to do well, not only in their target language learning but more generally in life since words play an important role in expressing one's feelings, emotions, and ideas to others during the act of communication in today's complex social and economic world (Allen, 1983). It follows that language teachers in general and S/FL teachers in particular should attribute great importance to teaching vocabulary in their classes.

The present paper goes beyond that and sees FL vocabulary teaching more than presenting new words. For sure, this may have its position but careful consideration should be given to other issues. For example, EFL students see lots of words in a language class, some of them are taught at an early stage and other are taught later and some are used often while others are not. At this level, one should ask why should instructors teach some words and not teach others? On which aspects should they base their EFL vocabulary selection decisions? Is there any way or source on which teachers may rely to encourage students to learn a given word rather than another? Why some students are successful at vocabulary learning and others are not? These questions are more or less difficult and challenging to foreign language syllabus or textbook designers, teachers and learners to answer for the large amount of the English vocabulary itself and the foreign language little exposure in the learning environment. Given the challenge involved in EFL vocabulary selection and instruction, the present paper strives to solve and answer some of the above queries by highlighting the way computer-corpus data can ease the difficulties involved in EFL vocabulary selection and instruction which is the scope of this paper, but prior to this the researcher tries to answer some questions, like: what is meant by computer-corpus data? And what can computer-corpus data tell EFL textbook designers or teachers about vocabulary?

## 2. Computer-corpus data

A computer-corpus is basically a collection of written or spoken texts which are stored in a computer. Written texts (newspapers and magazines) can be entered into a computer from a scanner, a CD, or the Internet. Spoken texts as conversations are recorded and then transcribed so as to feed them into the computer database. Eventually, it becomes possible to analyse the language with corpus software tools to see how people really speak or write (O' Keffee, McCarthy & Carter, 2007). According to many corpus-linguists (Sinclair J., 1991;

Hunston S., 2002; McCarthy M. J., 2004; O'Keeffe, A., M. J. McCarthy, & R. A. Carter, 2007) computer corpus data can tell us about:

### 2.1. Frequency

It means which words and expressions are most frequent and which are rare. The frequency of words is shown and ordered in a frequency list. This latter, tells us what word and phrases are used most often i.e. high frequency words; and which are less used i.e. low frequency words. As far as teachers are concerned, frequency lists will be useful for us to make choices about what vocabulary to teach and in what order to present it to our pupils. For example, when one makes a look at some corpora like Brown Corpus or Cambridge International Corpus he may notice that idioms use is rare. Thus, as a teacher, he will teach idioms later in the language programme. On the other hand, he can notice which items in a large vocabulary set, people talk most about in order to give them priority in teaching. These may include names of colours, pets, clothing and health problems. Based on a sample of 4.5 million words of spoken data from Cambridge International Corpus it has been noted that *The* is the first word in the rank frequency list for the whole corpus; the five most common verbs (apart from parts of the verbs be, have and do) are *said, know, see, get, and made*; the most common nouns are *time, people, way, years, work, government, man, day, work, and life*; the most common adverb is *so*. One can also see which words are more frequent than other similar or related words: *Yeah* is more frequent than *yes*; also *is* more frequent than *too*.

Frequency lists can also provide very useful statistics to help textbook writers present grammar items in the best way. If we take example of the verb *must* and consider the hundreds of uses of *must* in the Cambridge International Corpus (Spoken Corpus), the results reveal that, on average, only 5 % of all its uses are connected with *obligation* (e.g., you must have a visa to enter the United States). Another 5 % are in expressions, such as *I must admit* and *I must say*. Yet, the crushing majority of uses of *must* are to express *prediction*, such as: *that must have been nice; you must be hungry, etc.*

### 2.2. Differences in speaking and writing

As we saw in the previous section, corpus tools can give teachers information about how frequent a word is in different corpora, therefore, they can compare the frequency of vocabulary in newspapers, academic texts and conversations. For instance, the word *probably* is about five times more frequent in conversation than in newspapers and ten times more frequent in conversations than in academic texts. On the other side, *however* is eight times more frequent in newspapers than in conversation and over twenty times more frequent in academic texts than in conversation. Awareness of such differences helps teachers a great deal to decide on whether to insert and present vocabulary units in a written or spoken context.

### 2.3. Contexts of use

The corpus tells us about the situations in which people use certain vocabulary. Owing to this language source, it becomes possible to know whether an item of vocabulary is used by anyone and in whatever situation or particularly in more polite situations and relationships. Large corpora enable the possibility of presenting and using vocabulary appropriately for teachers and learners respectively. Examples are: formal and informal usage of expressions like *goodbye* and *see you*.

### 2.4. Collocations

The term *collocation* means a combination of two or more words often used together more frequently than would happen by chance. Collocation software enables us to find out all the words which are often combined

with others (see figure 1). Computer-corpus data is very helpful for searching for collocations of verbs like: *have, get, make, and do* i.e. *delexical verbs*. As their name implies these verbs do not have a lexical meaning of their own. They rather take their meaning from the words that they collocate or are used with. Providing collocational knowledge to our students, by means of those data, will not only lessen their word-choice, but make them sound more native-like, too.

**Make:** sure, difference, sense, decision, mistakes, decisions, money, mistake, reservations, copies and effort.

**Do:** anything, something, things, job, well, nothing, work, whatever, aerobics, gardening, homework and laundry.

(Adapted from McCarthy: 2004)

Fig. 1. Collocations of the words make and do

### 2.5. Strategic use of vocabulary

Strategic vocabulary means the words and expressions that are used to organize and manage discourse. Written texts are easily found in newspapers, books, on the Internet, etc. and then used by teachers as models for teaching words and expressions that writers use strategically to organize their texts. These include conjunctions like *along with* and *however*, which organize ideas within and across sentences and adverbs as *first, secondly*, etc., which list ideas within a paragraph or text. Yet, it is not easy to find words and expressions that speakers use to manage their different talks. So how can we find them? Corpus software can answer this need since it analyses different conversations and orders frequency lists of the strategic vocabulary speakers use to manage their speech. These include expressions which linguists call '*vague language*' as: *or something, and that kind of thing* and *and stuff* that speakers use often informally instead of specific words to refer to things, activities, or situations. Other examples of strategic vocabulary of conversations are discourse markers like '*anyway*' which speakers use with words like *so* and *well* to come back to the main point after an interruption. These strategies are not just extras or a plus in foreign language learning, but they are a must in creating "proper" dialogues and in creating "fine" relationships between speakers (Carter & McCarthy: 2006). Computer-corpus data, indeed, proves to have much to offer to EFL teachers and textbook writers especially while designing English vocabulary teaching materials that recently much attention is paid to as they pervade all four skill areas of listening, speaking, reading and writing.

### 3. Objectives of the research

From this research one aims to:

- Explore the extent of awareness of EFL teachers to the concept of computer corpus-data.
- Raise awareness on the effectiveness of computer-corpus data in the area of EFL vocabulary selection and instruction.
- Analyze the EFL teachers' attitudes on the application of computer-corpus analyses as a language source for EFL vocabulary selection and instruction.

### 4. Research questions

In fact, this research paper intends to answer problem area of this research by tackling the following questions:

- What language sources do EFL teachers usually rely on to teach about English vocabulary?
- How do EFL teachers qualify the existing EFL vocabulary selection sources and teaching materials?
- Are EFL teachers knowledgeable about computer-corpus data?

- What are teachers' perspectives regarding the application of computer- corpus data in EFL vocabulary selection and instruction?

## 5. Research methodology

### 5.1. Method

The choice of the method is mainly based on the problem of the investigation. In fact, the nature of the problem, the population and the looked-for data made the researcher decide on two methods. One chooses the descriptive and analytic methods because the work is mainly based on some definitions, mainly computer corpus-data and the description and the analysis of EFL teachers' attitudes on those data. That is, in order to report EFL teachers' perceptions on the impact that computer corpus-data has on English vocabulary selection and instruction, the researcher opts for the use of the aforesaid methods. Indeed, those methods enable one to adequately describe and analyze the collected data about the subject matter.

### 5.2. Sample population

To gather data about the EFL vocabulary teaching situation in general and teachers' attitudes on the application of computer corpus-data on English vocabulary selection and instruction; one needs a sample population of EFL teachers for the administering of a semi-structured interview. The population of teachers for this study is of 10 full-time teachers (6/10 are females and 4/10 males), who are charged with teaching English as a Foreign Language in the Department of Foreign Languages – Section of English – at the University of Tlemcen (Algeria). The research subjects were chosen at random because in random sampling, according to Brown (2001:72) "*each individual in the populations must have an equal chance of being selected*", which reduces the effect of bias and enhances objectivity.

### 5.3. Research instrument

Following the requirements of the present research, one makes use of the *semi-structured interview* to answer the aforementioned research questions and to supply the research subjects with the needed information about computer corpus-data if teachers show unfamiliarity about those data. Yet, in introducing the semi-structured interview, the investigator and the interviewees discussed some inevitable points about EFL vocabulary teaching; like EFL vocabulary teaching goals and importance to foreign language teaching and learning milieu. All teachers confess the importance of vocabulary teaching in the foreign language classroom. Some explained that vocabulary knowledge makes EFL learning more effective as words are used in both language perception and production and relate its inadequacy to the students' low level of English proficiency. Accordingly, Beglar and Hunt (2005:7) write "*vocabulary acquisition is a crucial, and in some senses, the central component in successful foreign language acquisition*". Others admitted that vocabulary is inevitable in teaching and that it should be presented, taught and recycled very often in and outside the classroom for vocabulary knowledge can ever be fully mastered.

To manage this interview, the researcher takes an interview schedule which lists the topics to be discussed out of beforehand determined questions. Yet, at the same time the investigator allows an elaboration in the questions and answers whenever needed. In collecting information, one writes down quick and abbreviated notes on the interviewees' answers as they code switch while speaking. Later on, the researcher summarizes the data gathered from each interviewee in order to organize and analyze the results in graphs.

## 6. Results and discussion

### 6.1. Results and discussion of the first research question

To see what language sources do EFL teachers use in order to teach about vocabulary, the following question was asked: Which language sources do you use to teach about the English vocabulary?

The results show that all teachers rely on the Oxford Advanced Learner's Dictionary and some vocabulary teaching books like, English Vocabulary in Use, as far as word knowledge and vocabulary activities are concerned. Yet, this query brought a hot debate between the interviewees so that two points of controversies rose up among the research informants. Some teachers (4/10) admitted that they hardly refer to vocabulary teaching books for the reason that their learners are adults enough to undertake their own vocabulary learning process. Other teachers (6/10), however, confessed on a planned vocabulary teaching is needed and added that vocabulary is inevitable in teaching and that it should be presented, taught and recycled very often in and outside the classroom for vocabulary knowledge can ever be fully mastered.

### 6.2. Results and discussion of the second research question

Subsequently, the informants were asked to answer the question: “How do you qualify the existing EFL vocabulary selection sources and teaching materials, if any?” to know whether or not they are satisfied with the EFL vocabulary sources and teaching materials they are currently using. The findings, in the graph below, show that the majority of teachers (6/10) are not satisfied by the English vocabulary teaching sources and materials they are currently using, unlike one teacher. The rest of the respondents show a neutral attitude as they choose to say ‘Average’. What is more is that this query sheds light on some EFL instructors' vocabulary teaching difficulties. Actually, those teachers who were not satisfied with the existing EFL vocabulary selection sources and teaching materials do relate their dissatisfaction to the problem of the outdated books in the library of the department and other times to the lack of vocabulary teaching books, to finally admit that they find the refuge in the Internet rather than the library to better their English teaching in general and that of vocabulary in particular.

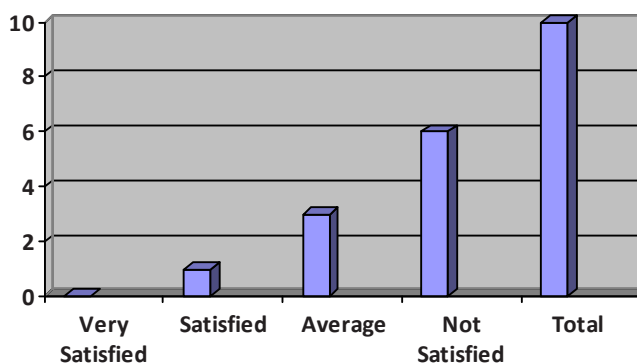


Fig. 2. Teachers' satisfactory level on EFL vocabulary teaching sources and materials

### 6.3. Results and discussion of the third research question

The third research question that the investigator endeavours to answer reads the following: Are EFL teachers knowledgeable about computer-corpus data? One aims from this question to introduce the notion of computer corpus-data to the research informants and to see whether or not teachers have prior knowledge about those data. In fact, once the researcher mentioned the idea of computer corpus-data and its investments in language analyses in general and vocabulary selection and teaching in particular, almost all research informants (9/10) reveal their unfamiliarity about those data (See Figure 3). Consequently, one has provided the sample of teachers with the relevant information about computer corpus-data in order to familiarize them and raise their awareness on the notion of those data. Furthermore, the supply of EFL teachers with the relevant information vis-à-vis computer corpus-data is fundamental since this made the collection of their attitudes on the application of those data as a source for EFL vocabulary selection and teaching possible.

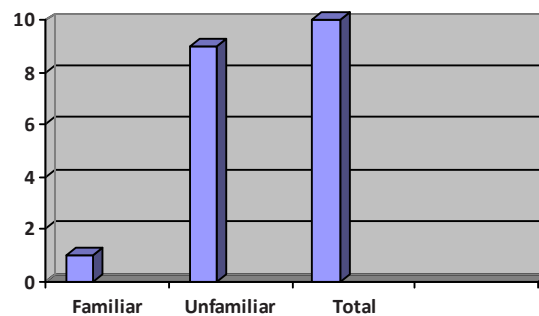


Fig. 3. EFL Teachers' Familiarity or Unfamiliarity with Computer Corpus-Data

### 6.4. Results and discussion of the forth research question

In an attempt to answer and gather data about the forth research question which enquires on EFL teachers' attitudes a propos the application of computer- corpus data in EFL vocabulary selection and instruction, the researcher had asked the research subjects to rank their perceptions on the statements below using the five Likert-type scale.

- Statement 1: Computer corpus-data is the 'richest' language source that syllabus designers and textbook writers may relay on for presenting both current and authentic language content and teaching materials as far as EFL learners are concerned.
- Statement 2: EFL vocabulary is best analysed, selected, presented and taught with the aid of computer corpus-data analyses.
- Statement 3: Raising awareness about computer corpus-data language analyses and recommending its application in EFL vocabulary teaching and learning situations may not only enhance EFL teachers and learners' interest in vocabulary teaching and learning but do also improve their proficiency level as a whole.

The findings show that the highest scores ranged between "Agree", "Strongly Agree", and "Agree" without leaving a single score to "Disagree" or "Strongly Disagree" as regards statements 1, 2, and 3 respectively. This



means that the research informants do very often show agreement on the above statements as illustrated in the graph below:

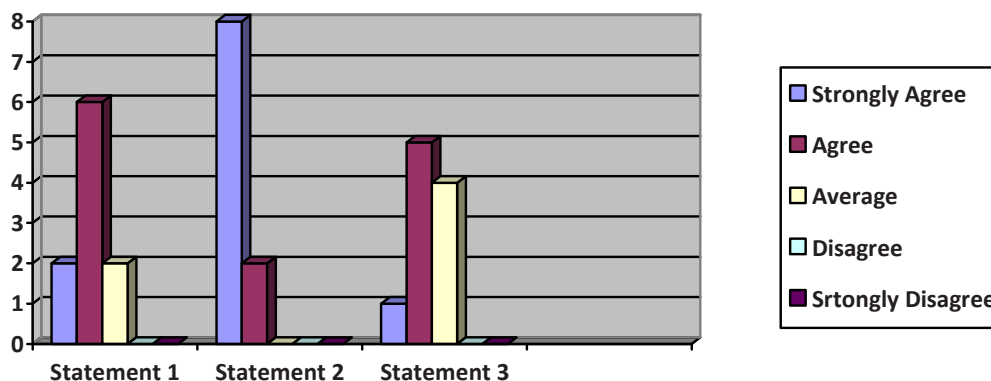


Fig. 4. EFL teachers' levels of Agreement and disagreement on statements 1, 2 and 3

While discussing the above statements, the research informants were conferring positive attitudes towards the application of computer corpus-data as a source for EFL vocabulary selection in particular and English teaching in general. They added that its use would bring *"new and light breath"* to EFL teachers and learners alike. What is more, the interviewees admitted that their learners would *"enjoy"* the EFL class more than ever for they will deal with a more authentic language that they expect to meet in the international mass media or receive while visiting English-speaking countries or while chatting in text or speech with English native speakers and non-native English speakers by means of social networks (Facebook, MSN, Skype and the like). In due course, the research subjects confess on the effectiveness of computer corpus-data and make recommendations on its application in EFL vocabulary selection and instruction.

## 7. Conclusions and Implications

There is no doubt that learning a foreign language is strongly based on the learning of its vocabulary because a widely productive vocabulary is essential in enhancing the learner's four language skills (listening, speaking, reading and writing). It follows that teachers are advised to plan their English vocabulary teaching and enforce its recycling and testing as word knowledge encompasses EFL learners' receptive and productive skills and to emphasize the importance of word knowledge and vocabulary teaching in general in their EFL classrooms by explaining its efficiency to foreign language learning.

Now it is clear that the corpus massive collection of texts and conversations, that were previously unavailable, have given language researchers, syllabus designers and instructors access to a wealth of information for both spoken and written English in order to determine the words and phrases needed to fulfil the EFL learners' vocabulary needs for conducting conversations and writing passages more effectively. From this, it is highly advocated for syllabus designers and teachers to apply computer corpus-data as a source for EFL vocabulary selection and instruction in particular and why not EFL teaching and learning in general.

From this research, however, it has been discovered that though EFL teachers confess the significance of vocabulary teaching in foreign language learning they do not have the tendency to implicitly teach about vocabulary on their classrooms and to update their language teaching sources and teaching materials in general



and that of vocabulary in particular since many of the research informants were unfamiliar with computer corpus data. And therefore, it is kindly requested for EFL teachers to update their teaching resources and materials for a better English teaching and learning.

## Acknowledgements

The author of this paper is thankful for the EFL teachers at the Department of Foreign Languages (in Tlemcen University) who participated in this study. I also thank the anonymous reviewers for their kind reading and comments.

## References

- Allen, V.F. (1983). *Techniques in teaching vocabulary*. Oxford: Oxford University Press.
- Beglar, D., & Hunt, A. (2005). Six principles for teaching foreign language vocabulary: A commentary on Laufer, Meara, and Nation's 'ten best ideas'. *The Language Teacher*, 29 (7), 7-10.
- Brown, T.D. (2001). *Using surveys in language programs*. Cambridge: Cambridge University Press, 2001.
- Carter, R. (1988). Vocabulary, Cloze, and Discourse. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 161 - 180). London: Longman.
- Celce-Murcia, M. (2001). *Teaching English as a second or foreign language* (3<sup>rd</sup> ed.). London, Heinle Heinle – Thomson Learning.
- Decarrico, J.S. (2001). Vocabulary learning and teaching. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (3<sup>rd</sup> ed., pp. 285 - 299). London: Heinle & Heinle Thomson Learning.
- Descamps, J. L. (1992). Towards classroom concordancing. In P. J. L. Arnaud, & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 167 - 181). Basingstoke, Hants: Macmillan.
- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Laufer, B. (1986). Possible changes in attitude towards vocabulary acquisition research. *IRAL* 24 (1), 69-75.
- Lewis, M. (1993). *The Lexical approach: The state of ELT and a way forward*. London: Language Teaching Publications.
- Lewis, M. (1997). *Pedagogical implications of the lexical approach in second language vocabulary acquisition: A Rationale for Pedagogy*. Cambridge: Cambridge University Press.
- McCarthy, M. (2004). *Touchstone: from corpus to course book*. Cambridge: Cambridge University Press.
- Nation, P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.
- Nattinger, J., & DeCarrico, J. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Nunan, D. (2001). Syllabus design. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (3<sup>rd</sup> ed., pp. 55-65). New York: Heinle & Heinle Thomson Learning.
- O'Keeffe, A., McCarthy, C., & Carter, R. (2007). *From corpus to classroom*. Cambridge: Cambridge University Press.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching: an anthology of current practice*. New York: Cambridge University Press.
- Sinclair, J., & Renouf, A. (1988). A lexical syllabus for language learning. In R. Carter, & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 140 - 158). London: Longman.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Swan, M. (2002). Seven bad reason for teaching grammar - and two good ones. In J. C. Richards, & W. A. Renandya (Eds.), *Methodology in language teaching: an anthology of current practice* (pp. 148 - 152). Cambridge: Cambridge University Press.
- Willis, D. (1990). *The Lexical syllabus*, London: Collins Cobuild.